Estudio de errores en la localización en espacios interiores, de dispositivos móviles por medio de tecnologías WiFi.

Gomez, Carlos Andrés. Perez, Antoni.
andres.gomez.ruiz@gmail.com, aperezn@uoc.edu
Universitat Oberta de Catalunya

Resumen—La localización en ambientes de interiores basada en dispositivos modernos de uso masivo, como los dispositivos móviles, es un área actual de investigación y desarrollo que ha empezado a generar productos para diferentes aplicaciones. Sin embargo, la estimación del error en estas mediciones, sus fuentes y la forma como se propaga, que evita tener mayor exactitud en la localización, aún presenta un área pendiente de estudio. Esta investigación toma una de las técnicas comunes de estudio de localización de dispositivos por redes inalámbricas *WiFi* para estudiar los diferentes errores que ella conlleva, sus fuentes y su propagación. Este trabajo también propone un nuevo enfoque y nueva metodología, que permite reducir el error en la localización de dispositivos WiFi, con base a la clasificación de las redes inalámbricas presentes en un escaneo, la identificación de muestras y datos fortuitos, la estabilización de muestreos de potencias de redes WiFi, y la implementación de un filtro de Kalman, para la predicción de medidas estables de estas potencias.

I. INTRODUCCIÓN

Trabajar con redes WiFi implica captar las señales de radiofrecuencia que básicamente poseen tres variables que contienen informaciones importantes y relativas a la localización: el ángulo en que la señal llega (Direction of arrival, DOA) para estimar la localización relativa de un dispositivo respecto al otro, el instante en que la señal llega (Time of arrival, TOA) con la cual se puede estimar el tiempo que tomó en propagarse la señal y así estimar la distancia que separa un dispositivo y otro, y la potencia con la que la señal es recibida (RSSI - Received Signal Strength Indication) que también sirve para estimar la distancia que separa un dispositivo y otro [1].

Los ambientes de interiores se caracterizan porque las ondas de radiofrecuencia se propagan por múltiples caminos debido a los fenómenos físicos asociados a estas y a la tecnología WiFi, como la difracción, absorción o reflexión de ondas electromagnéticas; la disposición y cantidad de antenas en los equipos transmisores, etc. Por otro lado, los cambios en las condiciones atmosféricas como la temperatura también puede afectar la propagación de las ondas de radio y su correspondiente potencia RSSI. Además, la frecuencia de 2.4GHz que ha sido ampliamente usada para el desarrollo de redes WiFi, es la frecuencia de resonancia del agua, así que las personas absorben ondas electromagnéticas en la banda de 2.4GHz [2]. Lo anterior implica una gran incertidumbre en las posibles predicciones de radiación en interiores y en una fuente de errores.

Los ambientes de interiores se caracterizan por tener gran cantidad de obstáculos para la comunicación directa entre los equipos, constituyendo escenarios de no linea de vista NLOS (Non Line Of Sight) [3]. Esta condición vuelve mucho más difícil e imprecisa la interpretación de los datos relativos a la localización, que llegan al dispositivo móvil, por lo anterior, la mayoría de esfuerzos se orientan a trabajar escenarios de obstrucción de la línea de vista.

En cuanto a las variables que contienen información relativa a la localización, el ángulo de recepción (DOA) y el tiempo de recepción (TOA), comparadas con la RSSI, presentan características de desempeño muy sensibles a la disponibilidad de linea de vista LOS; así que las técnicas de localización por TOA y DOA en escenarios de NLOS están sujetos a errores considerables. El desempeño de la RSSI es alterada medianamente por la falta de linea de vista LOS. Los escenarios NLOS llevan a efectos inesperados en relación a la

función potencia-distancia que pueden ser reducidos aplicando técnicas de filtrado a las mediciones. Así, muchas de las técnicas de identificación, mitigación y localización en escenarios NLOS han sido diseñadas basándose en la variable de RSSI. [4]

Otro de los problemas de la localización basada en el tiempo o ángulo de llegada (TOA y DOA) es la necesidad de hardware adicional que debe ser instalado sobre la red WiFi existente, así como la necesidad de un sistema especial de sincronismo de tiempos. Para el caso de la localización por DOA son necesarios arreglos especiales de antenas directivas de alto costo. Lo anterior, sumada la comentada necesidad de un escenario de línea de vista para el uso de TOA y DOA hace que no sean viables para la ubicación de dispositivos móviles en ambientes de interiores [5].

Otro aspecto importante que ha orientado las investigaciones de sistemas de localización de dispositivos móviles, es la necesidad de reutilizar la infraestructura existente y ampliamente difundida en espacios de interiores, como lo son las redes WiFi. Los puntos de acceso (AP Access Points) de las redes WiFi, como indicadores de localización, eliminan la necesidad de costos de infraestructura adicional [6].

Debido a la compleja naturaleza y comportamiento de la propagación de ondas de radiofrecuencia en espacios interiores, su difícil predicción y su comportamiento inestable, no es posible confiar que el nivel de potencia RSSI corresponda directamente con la distancia de separación entre los dispositivos de la red WiFi, condición necesaria para aplicar técnicas de localización como la triangulación; así que la técnica más ampliamente usada es la basada en el estudio de las recepción de la potencia de señal RSSI en el espacio interno que desea ser usado para la localización, proceso llamado '*fingerprinting*'. En la mayoría de los trabajos presentes en la literatura de localización WiFi según [1], [6], [8], [9], [10], [11] y [12], el proceso de localización se hace en dos etapas. En la primera etapa, o de calibración, se construye una base de datos que almacena la información de las mediciones de RSSI de las señales provenientes de los diferentes AP que se encuentran en el ambiente a estudiar. Estas mediciones se hacen con un dispositivo móvil que realiza observaciones en las diferentes ubicaciones físicas del ambiente. Esta base de datos se suele llamar *radio-map* o mapa de RSSI. En la segunda etapa, o de localización, el dispositivo a ser localizado toma medidas de RSSI de su posición y son comparadas con las posiciones grabadas en la base de datos de RSSI. A partir de esa comparación se realiza la localización del dispositivo [11]. El método que se usa para hacer la comparación de los datos en la estimativa de localización es lo que diferencia cada método desarrollado.

Descamps-vila [9] realiza un trabajo experimental en el que se parte del análisis euclidiano expuesto por el proyecto RADAR, para aplicarle técnicas de filtrado probabilístico a las mediciones de RSSI en la etapa de calibración del *fingerprinting*. Se propone el uso de filtrado de mediciones que no satisfagan un umbral dado de desviaciones estándar, con lo cual se logra medir la dispersión estadística de las muestras, descartando mediciones inestables que son las fuentes comunes de error. También se propone limitar el número de AP usados para la localización de un dispositivo dado, privilegiando las señales de AP que presenten mejor comportamiento en el análisis de dispersión estadística.

II. METODOLOGÍA DE LOCALIZACIÓN ESTUDIADA

Descamps-Vila, Perez-Navarro y Conesa [9] proponen realizar un filtro estadístico para estudiar la dispersión e inestabilidad de las mediciones de RSSI. El criterio usado por ellos es fijar un umbral superior en el análisis de desviación estándar hecho sobre las mediciones de RSSI de cada una de las señales de Access Points medibles en cada

ubicación del mapa de RSSI para la calibración del *fingerprinting*.

Descamps-Vila [9] propone el siguiente criterio:

$$si\,6_{i,j} \geq 3.0\,el\,valor\,lc_{i,j}\,será\,descartado.\,(1)$$

$$si\,6_{i,j} < 3.0\,el\,valor\,lc_{i,j}\,será\,usado\,para\,la\,calibración.\,(2)$$

donde σ es la desviación estándar, *j* representa cada AP, *i* representa el nodo de medición. lc_{ij} representa la muestra compleja de RSSI del respectivo AP en la ubicación dada:

$$lc_{ij} = \{j_{1,i}, j_{2,i}, \ldots j_{n,i}\} \quad (3)$$

lc_{ij} es un conjunto de valores de potencias RSSI de *n* cantidad de muestras en la posición *i*, cada una expresada en unidades de dBm.

Con lo anterior, se hace un muestreo de las señales WiFi presentes en un nodo de medición del mapa de RSSI para la calibración del *fingerprinting*, y se descartarán las señales de WiFi de las muestras que según el criterio de desviación estándar expuesto, sean más dispersas. Sobre las muestras aceptadas por el filtro, se calcula el promedio y se obtiene así un único valor de potencia RSSI para cada AP en la ubicación. Se obtiene un vector de potencias RSSI en donde el número de elementos representa la cantidad de AP en ese nodo de medición.

Con esta metodología se repite el análisis en cada uno de los puntos de medición hasta completar la construcción del mapa de RSSI, dado por una matriz de calibración:

$$Cal_Matr = lc_{ij}, i \in [0, s], j \in [0, k] \quad (4)$$

Donde *i* se refiere a cada nodo de medición, *j* se refiere a cada AP, *s* es la cantidad de nodos y *k* es la cantidad de AP.

Para la etapa de localización [9] propone una metodología en la que el usuario realiza mediciones de potencias RSSI de señales WiFi en un lugar dentro del espacio estudiado. A estas muestras se le aplica el criterio de desviación estándar de las ecuaciones 1 y 2. Se obtiene un vector de los promedios de las mediciones de RSSI de cada AP de la ecuación 3:

$$Lista_Medic = lp_j, j \in [0, m] \quad (5)$$

Donde *lp* es el valor promedio de las potencias RSSI de cada AP, *j* representa cada AP y *m* es el número total de AP que han pasado el filtro de desviación estándar en ese lugar particular.

Luego se usa el algoritmo de calculo de distancias euclidiana, a partir de las señales de RSSI, expuesta por Teuber [13] que permite hacer una aproximación de la diferencia de radio (li) entre muestra del usuario (P) y cada una de las posiciones muestreadas (N) en la calibración del *fingerprinting*:

$$l(P, N_i) = l_i = \sqrt{\sum_{j=1}^{r} (\bar{lc}_{ij} - \bar{lp}_j)^2} \quad (6)$$

donde lc_{ij} representan los datos de calibración del mapa de RSSI para cada ubicación. lp_j indica el vector de mediciones de RSSI del usuario ordenado por los AP presentes, *j* representa cada AP, e *i* representa el nodo de medición.

Para estimar la posición *P* de un dispositivo en particular respecto al sistema de coordenadas, Teuber [9] sugieren usar las medias ponderadas obtenidas por la ecuación 6 de los nodos de calibración (*q*) más cercanos. Los nodos de calibración *q* seleccionados son los que tienen menor valor de *li* según la ecuación 6; se seleccionaron los 3 nodos de calibración de menor distancia, así que q=3. Con este criterio se puede obtener las coordenadas:

$$x = \frac{1}{\sum_{i=1}^{q}(\frac{1}{l_i})} \cdot \sum_{i=1}^{q}(\frac{x_i}{l_i}) \quad (7)$$

$$y = \frac{1}{\sum_{i=1}^{q}(\frac{1}{l_i})} \cdot \sum_{i=1}^{q}(\frac{y_i}{l_i}) \quad (8)$$

$$z = \frac{1}{\sum_{i=1}^{q}(\frac{1}{l_i})} \cdot \sum_{i=1}^{q}(\frac{z_i}{l_i}) \quad (9)$$

La metodología propuesta por [9] permitió obtener resultados en la localización con resultados muy variables y errores en la precisión muy dispersos que van desde los 0.18m a los 5.01m de exactitud en alguna de las coordenadas *X* o *Y*.

III. Errores en la localización de dispositivos móviles por Redes WiFi

Luego de identificar la metodología de localización de dispositivos por redes WiFi, se realizó su implementación en un ambiente de interiores con el fin de medir nuevamente sus resultados, poder identificar fuentes de error y proponer mejoras a esa propuesta.

Se aplicó la metodología diseñada por [9] en un espacio de interiores según la figura 1.

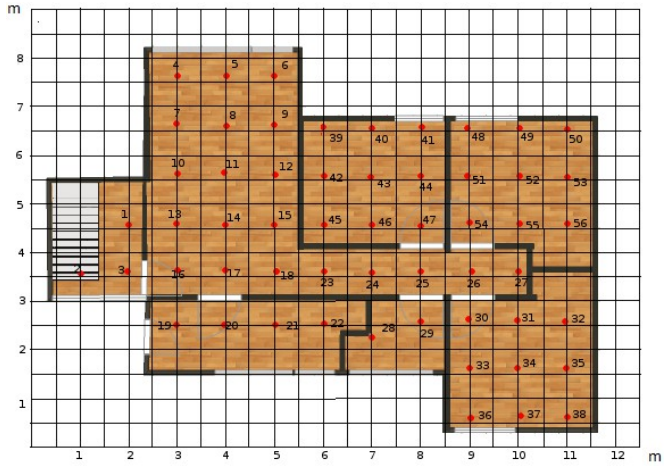

Fig 1. Plano de espacio estudiado y ubicación de los puntos de medición aplicados.

Para la toma de muestras de potencias de RSSI se realizó un aplicativo para teléfonos móviles con sistema operativo Android 4.0 que realiza escaneo de las redes WiFi durante 20 segundos; fueron tomadas 10 muestras de las redes WiFi, es decir, una muestra cada 2 segundos. Esta información es almacenada en una base de datos local. Las mediciones se realizaron colocando el teléfono en posición horizontal, siempre con la misma orientación y sobre una superficie de plástico de 1.2 metros de altura.

Se aplicó la metodología propuesta por [9] a la construcción del mapa de RSSI y a la localización de muestras de usuario tomadas en 10 ubicaciones diferentes dentro del espacio. Se obtienen los resultados resumidos en la tabla 1.

Tabla 1. Resultados de localización según metodología propuesta por [9]

	Posicion Real		Posicion Calculada con Desviacion estandar		Precisión (m)		Distancia de Error (de)
Ubicación	X	Y	X	Y	X	Y	Metros
1	4,5	5,2	3,43	5,99	1,07	0,79	1,33
2	5,5	2,5	6,17	2,68	0,67	0,18	0,70
3	7	3,7	5,71	3,66	1,29	0,04	1,29
4	6,5	5,7	5,41	3,28	1,09	2,42	2,66
5	8	4,7	7,32	4,98	0,68	0,28	0,74
6	8	2,5	7,01	4,23	0,99	1,73	1,99
7	9,5	5,2	7,30	5,00	2,20	0,20	2,21
8	10,5	4,7	10,00	5,69	0,50	0,99	1,11
9	10,5	2,2	10,04	1,24	0,46	0,96	1,06
10	9	1,7	9,66	1,06	0,66	0,64	0,92

La medición de la *distancia de error (de)* se halla mediante el cálculo de la distancia euclidiana entre la posición calculada y la posición real:

$$de(P) = \sqrt{(X_p - X_k)^2 + (Y_p - Y_k)^2} \quad (10)$$

Y la precisión en la medición es la diferencia entre la posición calculada y la real.

$$Precision_x = (X_p - X_k) \quad (11)$$
$$Precision_y = (Y_p - Y_k)$$

donde *P* hace referencia a cada una de las 10 posiciones estudiadas, *Xp* y *Yp* indican las coordenadas calculadas del dispositivo, y *Xk* y *Yk* son las coordenadas reales de la ubicación del dispositivo.

La distancia de error (*de*) mínima, la distancia máxima y la desviación estándar de los resultados de la tabla 1 se presenta en la tabla 2.

Tabla 2. Análisis de resultados de la tabla 1

Desv Est σ	0,66
Max	2,66
Min	0,70
Promedio	1,40

Lo anterior muestra que los resultados según la metodología propuesta por [9] es más o menos alta, con errores poco estables.

IV. ANÁLISIS DE ERROR

En el análisis de desviación estándar se encontró un problema asociado a las muestras que, aunque hacen parte de redes de potencia relativamente alta y altos porcentajes de aparición (son visibles en la mayoría de los muestreos), son filtradas ya que ocasionalmente alguna de las muestras tiene un valor atípico. Para ejemplificar lo anterior se presenta el caso de la tabla 3, que representa una situación bastante común.

Tabla 3. Ejemplo de muestreo con pocas muestras dispersas.

Nº muestra	valor muestra
1	-88
2	-82
3	-81
4	-75
5	-79
6	-76
7	-81
8	-75
9	-75
10	-78

En este caso la red inalámbrica se detecta con niveles relativamente altos, además de las 10 posibles muestras, se observa las 10 veces. Es decir, es una red que es bastante útil para ser usada en el análisis de localización, sin embargo el cálculo de desviación estándar (σ) es:

$$\sigma = 4.16$$

Con este cálculo esta red sería omitida por los filtros de las ecuaciones 1 y 2. Esto debido a la muestra número 1 de -88 dBm. Si dicha muestra se omitiera, es decir, solo se tuvieran en cuenta las muestras 2-10 el cálculo de desviación estándar sería

$$\sigma = 2.87$$

Con lo cual la red superaría el filtro de umbral por desviación estándar.

Estas variaciones que no representan un comportamiento altamente disperso no deben constituir factor para descalificar una red WiFi. Así, el criterio de filtrado por umbrales de desviación estándar presenta una fuente de error.

Otro error derivado de aplicar el filtro por umbrales de desviación estándar en la metodología de [9] es no incluir un criterio de frecuencia de las señales WiFi. Así el análisis de desviación estándar no tiene en cuenta cuantas veces fue visible un AP en un mismo muestreo, lo cual es un indicador directo de la estabilidad de este AP, ya que las señales que no son constantemente visibles en un nodo de medición son señales inestables y que no deben ser tomadas como buena referencia en la calibración del *fingerprinting*, ni en el proceso de localización, sin embargo estas señales pueden superar el filtro de desviación estándar. Para ejemplificar lo anterior se presenta el ejemplo de la tabla 4.

Tabla 4. Ejemplo de muestreo de redes con baja y alta frecuencia de aparición.

Nº Muestra	WLAN1 (dBm)	WLAN2 (dBm)
1	-88	-85
2	-88	-81
3	-85	-83
4	-85	-81
5	-	-80
6	-	-83
7	-	-83
8	-	-83
9	-	-83
10	-	-83
Desv Est σ	1,73	1,43

La tabla 4 ejemplifica cómo la aplicación simple del cálculo de desviación estándar en muestras de RSSI permite que señales inestables y poco confiables, sean tenidas en cuenta en todo el proceso de localización. La señal WLAN1 solo logró muestrearse un 40% de las ocasiones, mientras que la WLAN2 logró muestrearse el 100% de las ocasiones, sin embargo ambas redes superan el filtro de desviación estándar de las ecuaciones 1 y 2. La señal WLAN1 constituye una fuente impredecible de error por su aleatoriedad en la medición en un punto dado.

V. METODOLOGÍA PROPUESTA

Se apreció que en el escaneo de redes WiFi hay gran cantidad de redes lejanas que son perceptibles solo ciertas veces y constituyen unas medidas fortuitas que no deberían ser tomadas en cuenta. Para poder entender este comportamiento se graficaron los histogramas de la cantidad de apariciones de las redes WiFi en 10 escaneos, en todos los puntos de muestra para la calibración del *fingerprinting*. Algunos de los histogramas se muestran en la figura 2.

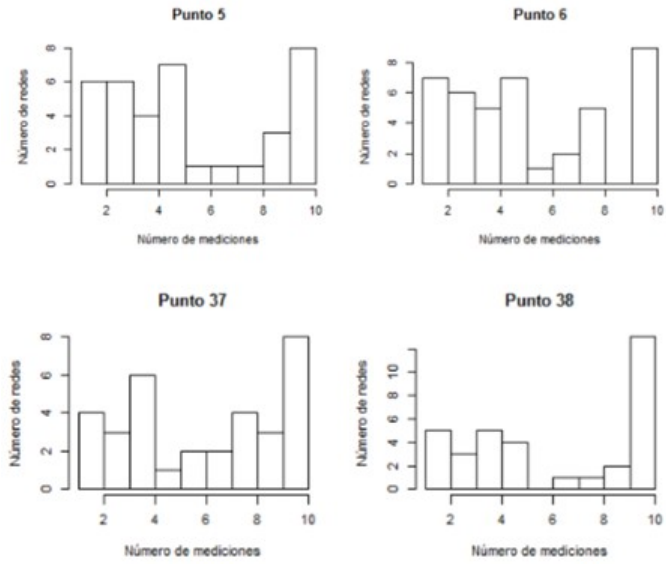

Fig 2. Histogramas de la frecuencia de aparición de redes WiFi en diversos puntos de medición.

La figura 2 permite ver que en cada muestra el comportamiento de las redes WiFi escaneadas es variable, sin embargo es general que las redes que tienen mayor cantidad de apariciones representen una cantidad importante de la totalidad de redes presentes. Así de las 6 mediciones ejemplificadas en la figura 2, la cantidad de redes WiFi que aparecieron al menos 7 de las 10 veces posibles (aparición superior de más del 60% de las veces, es decir su porcentaje de aparición es alto), se encuentran entre 10 y 16 redes.
En la mayoría de los puntos hay una gran cantidad de redes que son visibles solo unas pocas veces de entre las 10 posibles, es decir, tienen un porcentaje de aparición bajo.

La redes de bajo porcentaje de aparición corresponden a las redes distantes, que se perciben con baja potencia y comportamiento inestable. Por otro lado, las redes con porcentaje de aparición alto corresponden a las redes WiFi relativamente estables, cuya potencia de recepción no está en los umbrales bajos, y que presentan un escenario confiable de medición.

Así se implementó un primer criterio de filtrado, en el cual se omiten todas las redes cuyo porcentaje de aparición sea menor o igual al 60%, es decir, que de 10 posibles muestras estén presentes en 6 o menos.

Retomando la ecuación 3: $lc_{ij} = \{j1, j2, \ldots jn\}$

$$\text{Si } n >= 0.6 \cdot m \text{ la muestra es usada} \qquad (12)$$
$$\text{Si } n < 0.6 \cdot m \text{ la muestra no es usada}$$

donde n es la cantidad de muestras que hacen parte del conjunto lc_{ij} de un determinado AP (j) en la posición (i); y m representa la cantidad de muestras en la misma posición (i). Un ejemplo de la aplicación de este criterio al muestreo de un AP es:

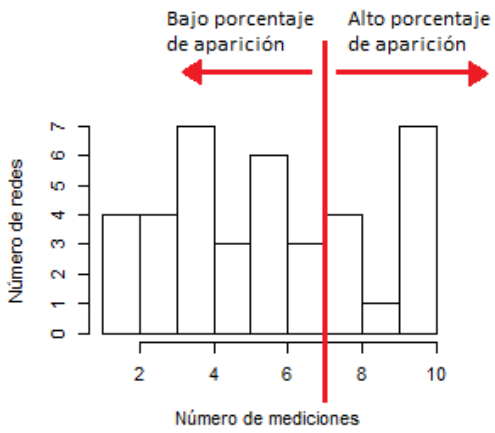

Fig 3. Definición de redes de bajo y alto porcentaje de aparición.

Con este primer criterio de filtrado se evita el uso de señales WiFi altamente inestables y poco confiables, en la etapa de calibración y en la etapa de localización.

Luego se propone un análisis estadístico para detectar las medidas atípicas en aquellas redes que han superado el filtro de la ecuación 12.
El diagrama de *boxplot* consiste en un rectángulo con los lados superior e inferior a los niveles de los cuartiles, una línea horizontal a nivel de la mediana, y los bigotes, de longitud 1,5 veces el rango intercuartíl, añadido en la parte superior e inferior. Los puntos fuera de estos límites se representan y son posibles valores atípicos [14]. Esta técnica permite identificar cada una de las muestras que representa valores atípicos, y así omitir cada una de esas muestras, para tener mediciones confiables, en vez de descartar *de facto* la red WiFi.

Se realizó análisis de *boxplot* a cada una de las redes de alto porcentaje de aparición en todos los puntos de medición de la calibración.

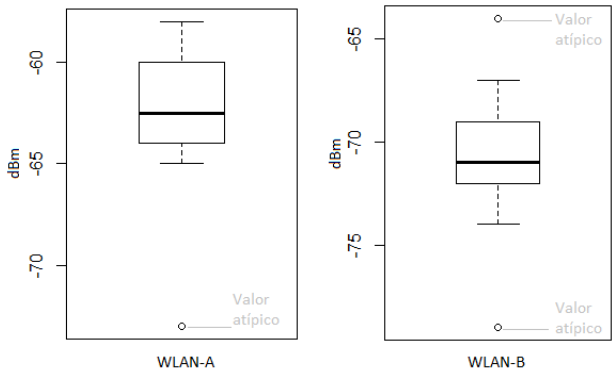

Fig 4. Ejemplo de análisis de *boxplot* para una red WiFi.

El ejemplo de la figura 4 muestra el análisis de *boxplot* de dos redes WiFi en un nodo de medición especifico. Allí se ve que las redes analizadas son susceptibles de tener mediciones atípicas fruto de las variaciones en el ambiente, los múltiples caminos en las propagaciones, y hasta la metodología de recopilación de las muestras. Sin embargo si se omiten esos valores atípicos, se obtienen muestras de gran información y que sí permiten caracterizar cada nodo de medición. El análisis de *boxplot* también permite entender la simetría de las muestras y la dispersión de las ellas.

Según lo anterior, el segundo filtro a aplicar es el de las muestras que son valores atípicos según la técnica de *boxplot*. No se omiten redes WiFi para su uso en el proceso de calibración o en el de localización, sino que se le hacen más estables.

Los dos procesos de filtraje se aplican a las muestras para la construcción del mapa de potencias RSSI en la etapa de calibración, y a las muestras para el proceso de localización del usuario. Así se construye un mapa de potencias más conciso en la etapa de calibración y se construye un vector de potencias más preciso para la localización del usuario.

Luego del proceso de filtraje de valores atípicos según la técnica de *boxplot*, se tiene el conjunto de mediciones de RSSI de cada AP resultante de la

ecuación 3 menos los valores atípicos identificados en el análisis de *boxplot*:

$$lb_{ij} = \{boxplot(j1, j2, \ldots, jn)\$out = FALSE\} \quad (13)$$

lb_{ij} es un conjunto de valores de potencias RSSI de *n* cantidad de muestras en la posición *i;* cada una expresada en unidades de dBm, donde se han sustraído los valores RSSI atípicos.

Para la construcción del vector de mediciones RSSI no se usa el cálculo de promedio de potencias de un mismo AP en una misma ubicación, como es propuesto por Descamps-Vila, Perez-Navarro y Conesa (2014), sino que se aplica el filtro de Kalman que es es una poderosa herramienta para reducir el ruido en las mediciones y poder predecir una medición estable. El filtro de Kalman es esencialmente un conjunto de ecuaciones que implementa un estimador del tipo predictor – corrector que es óptimo para minimizar la covarianza del error [16].

El filtro de Kalman para un modelo estático, como es el caso de una medición individual de la potencias RSSI de una AP en una ubicación específica, reduce las ecuaciones ya que no son necesarias algunas variables usadas en modelos dinámicos. Las ecuaciones del filtro de Kalman en la etapa de predicción para el modelo estático se representan en las ecuaciones 14 y 15.

$$\mu_{pt} = \mu_{t-1} \quad (14)$$

$$\Sigma_{pt} = \Sigma_{t-1} + R_t \quad (15)$$

Las ecuaciones del filtro de Kalman en la etapa de actualización de medidas se representan en las ecuaciones 16, 17 y 18.

$$K_t = \Sigma_{pt}(\Sigma_{pt} + Q_t)^{-1} \quad (16)$$

$$\mu_t = \mu_{pt} + K_t(z_t - \mu_{pt}) \quad (17)$$

$$\Sigma_t = (I - K_t)\Sigma_{pt} \quad (18)$$

Donde:

- μ es el estado estimado.
- Σ corresponde a la varianza de estado(por ejemplo el error en la estimación)
- *R* corresponde a la varianza del proceso (por ejemplo el error debido al proceso).
- *Z* son las mediciones.
- *K* es la ganancia de Kalman
- *Q* correponde a la varianza de medición (por ejemplo errores en la medición).

Los subindices indican: *t* tiempo actual, *t-1* periodo previo de tiempo, *pt* tiempos intermedios (Thrun, Burgard & Fox, 2000.).

La aplicación de las ecuaciones del filtro de Kalman crean un ciclo en el que es necesario asumir unas condiciones iniciales para calcular los primeros valores de las ecuaciones predictivas, y empezar la iteración de valores teniendo en cuenta que el estado previo será la entrada del estado actual. Se asume el primer valor de estado estimado μ de -93, debido a que para potencias RSSI este valor es el que representa la sensibilidad de la mayoría de tarjetas WiFi (-93 dBm) (Kaemarungsi & Krishnamurthy, 2004); la aplicación del Filtro de Kalman para un modelo estático necesita representar una varianza muy alta para lograr la convergencia del filtro con rapidez, así se representa una varianza de estado Σ de 1000. También se asume que el modelo adoptado es preciso, así que el ruido del sistema (*R*) es pequeño, 0.1. Finalmente se asume un valor de -5 (dB) para el ruido en la medición (*Q*) ya que los resultados de Prasithsangaree, P., et al (2002) indican que hay variaciones de hasta -5 dB en mediciones por factores externos a la red WiFi.

Se aplican estas ecuaciones, según la iteración propuesta por [15], a cada uno de los conjuntos de mediciones filtradas por las ecuaciones 12 y 13. Estas mediciones se usarán en la ecuación 17 de actualización de medidas.

La figura 5 ejemplifica la aplicación de las ecuaciones del 14 a 18 a mediciones de RSSI.

	Predict		Measurement and Update			
t	x_t|t-1	p_t|t-1	y_t	K_t	x_t|t	p_t|t
0	N/A	N/A	-93	N/A	**-93,0000**	1000,0000
1	-93,0000	1000,1000	-73	1,0050	**-72,8995**	-5,0251
2	-72,8995	-4,9251	-67	0,4962	**-69,9720**	-2,4811
3	-69,9720	-2,3811	-74	0,3226	**-71,2714**	-1,6130
4	-71,2714	-1,5130	-73	0,2323	**-71,6730**	-1,1615
5	-71,6730	-1,0615	-77	0,1751	**-72,6059**	-0,8756
6	-72,6059	-0,7756	-81	0,1343	**-73,7331**	-0,6715
7	-73,7331	-0,5715	-77	0,1026	**-74,0682**	-0,5128
8	-74,0682	-0,4128	-73	0,0763	**-73,9867**	-0,3814
9	-73,9867	-0,2814	-74	0,0533	**-73,9874**	-0,2664
10	-73,9874	-0,1664	-71	0,0322	**-73,8912**	-0,1610

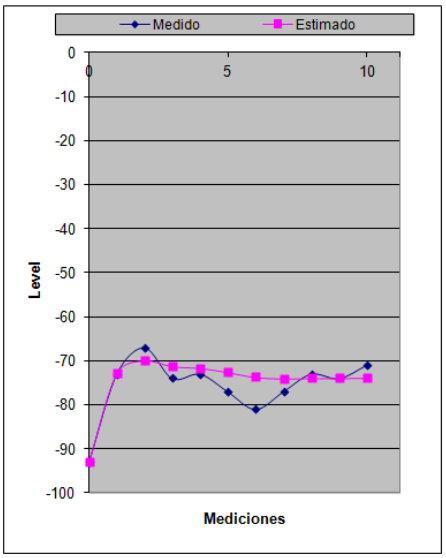

Fig. 5. Ejemplo de aplicación del filtro de Kalman para el cálculo de los valores de los vectores de RSSI en las mediciones de dos AP en un lugar específico.

En la Fig 5 la curva azul representa la gráfica de los valores medidos (Zt) y la curva rosa es la estimación de valores según el Filtro de Kalman (μ_t). Para construir los vectores de medidas y el mapa de RSSI se toma el último valor (μ_t más estable) estimado por el filtro de Kalman. Los vectores resultantes son similares en forma a los de la ecuación 5.

$$Lista_Medic = lk_j, j \in [0, m] \quad (19)$$

Donde *lk* es el último valor estimado (μ_t) del filtro de Kalman sobre las potencias RSSI, *i* representa cada AP y *m* es el número total de AP que han sido filtrados por las ecuaciones 12 y 13, en ese lugar particular.

Con estos vectores se construyó el Mapa de RSSI, a través de la matriz de calibración, como lo describe la ecuación 4. Para la construcción de los vectores de las muestras de los dispositivos a localizar se aplicó la misma metodología propuesta, y dicho vector tiene la misma forma descrita por la ecuación 19.

En el proceso de localización se usó la misma técnica usada por Laia Descamps-Vila, es decir la medición de distancias euclidianas para hallar las coordenadas *X* y *Y* del usuario, como para calcular el error en la localización (distancia de error); ecuaciones 6, a 10. Se usaron las mismas muestras usadas para el proceso de localización basada en filtraje por umbrales por cálculo de desviación estándar y cálculos de promedios, para aplicarlas en el filtraje por análisis por *boxplot* y aplicación del filtro de Kalman, y se obtienen los resultados de la tabla 5.

Tabla 5. Resultados de localización con metodología propuesta.

	Posicion Real		Posicion Calculada con Boxplot y Filtro Kalman		Precisión (m)		Distancia de Error (de)
Ubicación	X	Y	X	Y	X	Y	Metros
1	4,5	5,2	4,44	4,96	0,06	0,24	0,24
2	5,5	2,5	5,66	2,92	0,16	0,42	0,45
3	7	3,7	6,09	3,34	0,91	0,36	0,98
4	6,5	5,7	7,01	5,35	0,51	0,35	0,62
5	8	4,7	7,01	5,35	0,99	0,65	1,19
6	8	2,5	8,72	2,85	0,72	0,35	0,80
7	9,5	5,2	9,33	4,04	0,17	1,16	1,17
8	10,5	4,7	9,04	3,53	1,46	1,17	1,87
9	10,5	2,2	10,36	1,69	0,14	0,51	0,52
10	9	1,7	10,22	1,65	1,22	0,05	1,22

La Tabla 5 presenta los resultados en la muestra estructura de la Tabla 1 (resultados de la metodología de Laia Descamps-Vila) para su comparación.

Tabla 6. Análisis de resultados de la tabla 5.

	Distancia de Error (de) con Desviacion estandar	Distancia de Error (de) con Boxplot y FiltroK
Desv Est σ	**0,66**	**0,48**
Max	2,66 m	1,87 m
Min	0,70 m	0,24 m
Promedio	1,40 m	0,91 m

	Reducción de la Distancia de Error (de) en Metodología propuesta vs metodología de Desviación estándar
Desv Est σ	27,8%
Max	29,5%
Min	65,1%
Promedio	35,2%

La tabla 6 muestra como, con los mismos valores de análisis, la metodología propuesta basada en el filtrado por frecuencia de aparición del AP, la aplicación de la técnica de *boxplot* para descartar valores atípicos y la aplicación del filtro de Kalman para obtener un valor característico de potencia RSSI de cada AP, permite la reducción del error absoluto (distancia de error) en la localización de dispositivos WiFi en interiores; respecto a los resultados logrados con la metodología propuesta por Descamps-Vila, Perez-Navarro y Conesa (2014).

De la tabla 6 se deduce que la metodología propuesta tiene una exactitud en la localización, representada por la distancia de error, de entre 0.2m y 1.87m, y en promedio este error está por debajo de 1m. Los resultados reportados por otras experiencias encontradas en la literatura, y que se asemejan a la metodología estudiada se presenta en la tabla 7, de la cual se deduce que el error resultado es menor que en todas ellas.

Tabla 7. Resultados de distancia de error en otros sistemas de localización.

Sistema	Distancia de error
RADAR. [7]	2-3m
WaveLAN. [13]	3m
Horus. [17]	2m

VI. CONCLUSIONES

Se tomó la metodología de localización de dispositivos móviles por redes WiFi propuesta por Descamps-Vila [9] para estudiar las posibles fuentes de error en los cálculos de la posición, a partir de allí se propone una metodología mejorada que permite corregir las fuentes de error detectadas, y se obtienen mejores resultados en la exactitud de la localización de dispositivos.

Se detectaron problemas en la metodología de referencia que introducen errores que son propagados a lo largo de toda ésta, y dan como resultado distancias de error mayores en los resultados de localización. Dentro de estos errores están el no contar la frecuencia de aparición de las señales WiFi en los muestreos como elemento de importancia, y los efectos de aplicar umbrales por desviación estándar a muestras que tienen comportamientos impredecibles, asociados a situaciones externas o del ambiente.

La frecuencia de aparición de una red WiFi en el muestreo de redes, es un parámetro muy importante que refleja la estabilidad y confiabilidad de la presencia de una red en un espacio determinado. Clasificar las redes muestreadas en redes de bajo o alto porcentaje de aparición, permite reducir los riesgos de tratar con datos inestables.

El proceso de escaneo de señales de potencia RSSI en redes WiFi implica un escenario de gran incertidumbre y medición de muestras fortuitas, que son localizables por medio de la técnica estadística de *boxplot.* Esta técnica permite corregir el muestreo de señales WiFi, que de lo contrario serían altamente dispersas.

El proceso de filtraje de redes con bajo porcentaje de aparición, el filtro de muestras atípicas por *boxplot* y la aplicación del filtro de Kalman, permiten la estabilización de los muestreos y una mejor caracterización del espacio mapeado. Lo anterior resulta en una reducción de la distancia de error en la localización del 35,2% en promedio, en relación a los resultados obtenidos por la

metodología de Descamps-Vila, Perez-Navarro y Conesa (2014). El promedio de la distancia de error obtenido está por debajo de 1 metro.

El cálculo de desviación estándar (tabla 6) de estos resultados permite ver que tan precisos son los resultados de cada técnica analizada. La técnica propuesta obtiene una menor desviación, lo que indica que sus resultados son más precisos y su error es menos disperso.

La aplicación de la técnica de control estadístico de ruido en muestreos, filtro de Kalman, permite una mejor aproximación de un valor característico de un conjunto de muestras y así reducir errores producidos por la presencia de muestras con altos niveles de ruido.

La metodología propuesta presenta una reducción de, en promedio, 35% de la distancia de error asociada a la distancia euclidiana entre la posición calculada y al posición real; lo anterior se logró luego de filtrar las redes más inestables en el escaneo WiFi, y de estabilizar las mediciones y cálculos por medio del filtro de las medidas atípicas y la aplicación de un filtro de Kalman para reducir el ruido de las mediciones de potencias de RSSI.

VII. FUTUROS TRABAJOS

La metodología propuesta, que permite la reducción de la distancia de error en la localización, está diseñada para mediciones estáticas, es decir, donde el dispositivo a ser localizado no está en movimiento. Esta metodología debe ser acondicionada para lograr la localización de un dispositivo que se mueve a través del espacio.

La metodología se ha basado en mediciones hechas en una posición estática y sin intervención humana. Así, es necesario extender estudios que permitan localizar y minimizar errores asociados a la interacción humana, como podría ser el cambio en el plano vertical, cambios en la orientación del equipo, y la cercanía a cuerpos de agua absorbentes de radiaciones en la frecuencia de 2.4 Ghz.

REFERENCIAS

[1] Nuaimi, K. Al, & Kamel H. (2011). A Survey of Indoor Positioning Systems and Algorithms.

[2] Ladd, A. M., Bekris, K. E., Rudys, A., Kavraki, L. E., Wallach, D. S., & Marceau, G. (2002). Robotics-based location sensing using wireless ethernet. *Proceedings of the 8th Annual International Conference on Mobile Computing and Networking - MobiCom '02*, 227. doi:10.1145/570673.570674

[3] Kaemarungsi, K., & Krishnamurthy, P. (2004). Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting. The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, MOBIQUITOUS, pp 14-23

[4] Zekavat et al. (2011). *Handbook of Position Location*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

[5] Atia, M. M., Korenberg, M. J., & Noureldin, a. (2012). Particle-Filter-Based WiFi-Aided Reduced Inertial Sensors Navigation System for Indoor and GPS-Denied Environments. *International Journal of Navigation and Observation*, *2012*, 1–12. doi:10.1155/2012/753206

[6] Lee, J., Yoon, C., Park, H., & So, J. (2013). Analysis of Location Estimation Algorithms for Wifi Fingerprint-based Indoor Localization.

[7] Bahl, P., & Padmanabhan, V. N. (2000). RADAR: an in-building RF-based user location and tracking system. *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, *2*, 775–784. doi:10.1109/INFCOM.2000.832252

[8] Kaemarungsi, K., & Krishnamurthy, P. (2004). Modeling of Indoor Positioning Systems Based on Location Fingerprinting.

[9] Descamps-vila, Perez-Navarro, & Conesa, J. (2014). RSS and sensor fusion algorithms for indoor location systems on smartphones. Connecting a Digital Europe Through Location and Place. Lecture Notes in Geoinformation and Cartography 2014, pp 197-214

[10] Martin, E., & Friedland, G. (1947). Precise Indoor

Localization Using Smart Phones.

[11] Moura, I. (2007). WBLS : um sistema de localização de dispositivos móveis em redes Wi-Fi.

[12] Prasithsangaree, P., Krishnamurthy, P., Chrysanthis, P. K., & Program, T. (2002). On indoor position location with wireless LANs. 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications,vol. 2, pp. 720–724.

[13] Teuber A, Eissfeller B (2006) WLAN indoor positioning based on Euclidean distances and fuzzy logic. Proceedings of the 3rd Workshop on positioning, navigation and comunication pp 159–168

[14] Tukey, J.W. (1970). "Exploratory data analysis", Vol. 1, Ch. 5. Addison-Wesley Publishing Co. ISBN 978-0201076165.

[15] Thrun, P., Burgard, W., & Fox, D., (2000) Probabilistic Robotics. MIT ISBN: 9780262201629

[16] Welch, G. & Bishop, G. (2001) An Introduction to the Kalman Filter. En: Proceeding of SIGGRAPH 2001. ACM Press

[17] Youssef, M., Agrawala, A. K. (2004). ''Handling samples correlation in theHorus system,'' IEEE INFOCOM 2004, Hong Kong, vol. 2, pp. 1023-- 1031, Mar. 2004